

高等学校入学英语考 试(MET)答题卷

姓 名

考号

- K I 1 (A) (B) (C) (D) 6 (A) (B) (C) (D) K II 11 (A) (B) (C) (D)
 2 (A) (B) (C) (D) 7 (A) (B) (C) (D) 完 形 12 (A) (B) (C) (D)
 3 (A) (B) (C) (D) 8 (A) (B) (C) (D) 填 空 13 (A) (B) (C) (D)
 4 (A) (B) (C) (D) 9 (A) (B) (C) (D) 14 (A) (B) (C) (D)
 5 (A) (B) (C) (D) 10 (A) (B) (C) (D) 15 (A) (B) (C) (D)

--	--	--	--	--	--	--	--	--	--

- 0 (0) (1) (2) (3) (4) (5) (6) (7) (8) (9)
 1 (0) (1) (2) (3) (4) (5) (6) (7) (8) (9)
 2 (0) (1) (2) (3) (4) (5) (6) (7) (8) (9)
 3 (0) (1) (2) (3) (4) (5) (6) (7) (8) (9)
 4 (0) (1) (2) (3) (4) (5) (6) (7) (8) (9)
 5 (0) (1) (2) (3) (4) (5) (6) (7) (8) (9)
 6 (0) (1) (2) (3) (4) (5) (6) (7) (8) (9)
 7 (0) (1) (2) (3) (4) (5) (6) (7) (8) (9)
 8 (0) (1) (2) (3) (4) (5) (6) (7) (8) (9)
 9 (0) (1) (2) (3) (4) (5) (6) (7) (8) (9)

- K I 31 (A) (B) (C) (D) 36 (A) (B) (C) (D) 41 (A) (B) (C) (D)
 32 (A) (B) (C) (D) 37 (A) (B) (C) (D) 42 (A) (B) (C) (D)
 33 (A) (B) (C) (D) 38 (A) (B) (C) (D) 43 (A) (B) (C) (D)
 34 (A) (B) (C) (D) 39 (A) (B) (C) (D) 44 (A) (B) (C) (D)
 35 (A) (B) (C) (D) 40 (A) (B) (C) (D) 45 (A) (B) (C) (D)
 U I 61 (A) (B) (C) (D) 66 (A) (B) (C) (D) 71 (A) (B) (C) (D)
 62 (A) (B) (C) (D) 67 (A) (B) (C) (D) 72 (A) (B) (C) (D)
 63 (A) (B) (C) (D) 68 (A) (B) (C) (D) 73 (A) (B) (C) (D)
 64 (A) (B) (C) (D) 69 (A) (B) (C) (D) 74 (A) (B) (C) (D)

• 桂诗春

4.7

准化考试—理论、原则与方法

• 广东高等教育出版社

标准化考试——理论、原则与方法

桂诗春编著

广东高等教育出版社

标准化考试——理论、原则与方法

桂诗春 编著

广东高等教育出版社出版

广东省新华书店发行

广州红旗印刷厂印刷

787×1092毫米32开本 6.375印张 125千字

1986年10月第1版 1986年10月第1次印刷

印数 1—10,000册

统一书号7343·20 定价1.25元

前 言

呈现在读者面前的这本小书是一篇“急就章”，是在教学与行政工作之余，用不到三个月的时间赶出来的，疏漏错误之处在所难免，敬希国内名家不吝指正。

编这本书的起因是：

一、在当前我国考试制度改革中，实现标准化考试的问题被提出来了。什么是标准化考试？标准化考试在考试现代化中占什么位置？标准化考试是一种什么性质的考试？标准化考试有些什么特点？怎样编制和组织标准化考试？标准化考试有些什么利弊？怎样评估一个标准化考试？怎样通过标准化考试去促进教学？这些问题受到考试的组织人员、命题人员、使用考试的学校和单位、广大教师、考生、考生家长的普遍关注。1979年以来，笔者先后参加了原教育部所组织的英语水平考试(The English Proficiency Test, 简称EPT)，广东省所组织的高等学校入学英语考试(The Matriculation English Test, 简称MET)的设计和命题工作。这两个考试都是按标准化考试要求建立的；虽不算成熟，但终归使我国标准化考试的研究从理论上的探讨发展到实际上的试验，迈出可贵的一步。近年来，笔者曾应邀在全国各地作标准化考试的专题报告。最近的一次是在一九八六年初对广东省准备进行标准化考试试验的科目命题组成员和省、地区招生办公室负责人讲的，当时广州外语音像教材出版社曾作现场录象，并制成药五小时的录象带——《标准化考试的设计》公

开发行。但是，标准化考试是一项技术性很强的系统工程，在几个小时的讲话里，很多问题都语焉不详，遂有将讲话内容编辑成书的动议。因此，这本小书可以说是录象带的补充与完善。因为写书不受时间限制，所以在书中增加了不少材料。

二、由于标准化考试试验的开展，考生中产生一种要求了解考试、适应考试的迫切心情。为了迎合这种心情，各种“标准化试题集”、应试小册子一时充斥市场。由于这些书籍的编者大都自己也不懂得什么是标准化考试（如果懂得，也就不会编这类东西去贻害学生），而且大都是粗制滥造，于考生有害无益。这些书籍不但对考生全面复习功课，提高知识和能力水平起了严重干扰作用，而且也影响了标准化考试试验的顺利开展，甚至起了鱼目混珠的作用。有些好心人以为这些“试题集”就是标准化考试，把它作为标准化考试的弊端而非议。其实从事标准化考试试验的同志对这些现象更为深恶痛绝。所以编这本书的另一目的是正视听，减少误解，使标准化考试试验得以健康发展！

所以，这本书是写给社会上所有关心标准化考试的人看的。书中有些章节专业性较强，因为在关心标准化考试的人中有不少是考试的组织、管理人员、命题人员、统计分析人员，他们想知道标准化考试是怎样实现的。对这些章节，普通读者可以略而不看。本书也可作为大学和师范学院教育测量学课程的辅助读物。为了适应不同读者的情况，文字力求通俗易懂，并尽量辅以例子，对所使用的公式只介绍其用法，而不作推导。

中国是最早推行考试的国家之一，经验教训多，但因袭也重。标准化考试的一整套做法大多数是引进的，但在本书

所介绍的也并非生吞活剥地照搬，而是经过咀嚼消化，大多数是实践过，并证明是有效的。例如书中所介绍的试测和考试的检验均由广州外国语学院编制成电子计算机专用软件包，并已推广到全国数十所大专院校，使用多年。

对外国的东西，我们赞成鲁迅先生说过的“拿来主义”：“运用脑髓，放出眼光，自己来拿”。“运用脑髓”就是开动脑筋，分析批判，对外国东西既不一概盲从，也不一概排斥。在中国，“全盘西化”和闭关自守都吃过大亏。

“放出眼光”就是从实现四个现代化的目标出发，高瞻远瞩。新生事物的诞生与成长总不会一帆风顺，只要看准了方向，就要有胆识、有勇气，排除万难去争取胜利。“自己来拿”就是以我为主；借鉴外国，无非为了发展自己。“没有拿来的，人不能自成为新人。”这是鲁迅先生在文章中最后的呼吁。鲁迅先生的几句话正是我们对待标准化考试试验的基本原则：从实现我国四个现代化的高度出发，借鉴外国的先进经验，建立适合我国国情的考试制度。但是我们开始时要先有所引进，然后在使用过程中才有所吸收、有所改进、有所扬弃，也就是鲁迅先生所说的，“总之我们要拿来。我们要或使用，或存放，或毁灭。”（《鲁迅全集·第六卷》，31～33页）我们所不赞成的是那种以中国国情为借口而拒绝向外国先进经验学习的闭关自守的态度，例如对标准化考试是什么一回事还没有弄清楚，就一味加以反对。所以，在鲁迅先生所说的几条以外，不妨补充一条：“实践是检验真理的唯一标准”。对标准化考试（包括它所使用的试题）的评估必须从经验性的讨论转到实验性的检验。一个考试之好坏决定于它实际使用的客观效果，而不在于人们的主观的评论。例如，考试结束后我们都会听到各种议论：这次考试较

易（或较难），那些题目出得好（或出得不好）。这些议论大都是经验性的，或是根据局部的观察而得出的结论，亦有参考价值。但是考试是否真的较易（或较难），题目是否真的出得好（或出得不好），还必须看实际统计出来的数据，让事实说话。这需要：第一、大家要懂得一点考试中使用的统计分析方法。不一定要懂得如何具体运算，但起码要懂得怎样去解释各种统计出来的指标，就等于医生不一定懂得怎样去验血，但却要懂得看验血报告，然后才去分析病情。第二、主持考试的机构要定期地、客观地公布这些统计指标。

作为一种考试制度，标准化考试并非完美无缺的；作为一种制度去实施，问题和缺点还会更多。但这是前进中的问题和缺点。我们相信，经过广泛的实践和认真的总结，我们必将能建立起一套符合我国实际的现代化的考试制度！

桂诗春
1986年清明节

目 录

前 言

第一章 考试的现代化

- §1.1 考试中的弊端…………… (3)
- §1.2 制度化…………… (4)
- §1.3 专业化…………… (5)
- §1.4 标准化…………… (8)
- §1.5 电脑化…………… (10)

第二章 什么是标准化考试

- §2.1 从用途上区分…………… (14)
- §2.2 从反映分数的方法上区分…………… (17)
- §2.3 从考试的方法上区分…………… (18)
- §2.4 从考试的方式上区分…………… (19)
- §2.5 从考试的要求上区分…………… (20)
- §2.6 从考试的规模上区分…………… (21)
- §2.7 标准化考试的要求…………… (22)
- §2.8 几个应该注意的问题…………… (23)

第三章 标准化试卷的设计

- §3.1 标准化考试的实现流程…………… (27)
- §3.2 考试类型的决定…………… (28)
- §3.3 考试内容的决定…………… (30)
- §3.4 考试题型的决定…………… (33)
- §3.5 克服主观性试题的主观因素…………… (37)
- §3.6 客观性试题的编写…………… (44)
- §3.7 怎样对待客观性试题的猜测…………… (61)

§3.8	一份试卷应有多少选择题才算合适	(65)
§3.9	试卷比重的决定	(66)
§3.10	双向细目表的制订	(70)
§3.11	制定考试大纲	(72)
第四章	试测	
§4.1	试测的目的	(76)
§4.2	试测的过程	(77)
§4.3	组织试测	(77)
§4.4	题目分析	(81)
§4.5	题库的建立	(101)
§4.6	拼题	(104)
第五章	考试的实施	
§5.1	考卷的准备	(106)
§5.2	考场的组织	(110)
§5.3	考卷的机器评阅	(113)
第六章	常模的建立	
§6.1	分数的分布	(119)
§6.2	信度	(128)
§6.3	效度	(137)
§6.4	考试的分析报告	(155)
§6.5	编制百分位表	(156)
§6.6	标准分	(159)
§6.7	分数的转换	(160)
§6.8	对主观性试题分数的调整	(168)
§6.9	考试手册的编写	(173)
附录 I	高等学校入学英语考试大纲	(176)
附录 II	正态曲线的面积和纵线表	(181)
附录 III	本书所用考试术语的汉英对照表	(191)

第一章 考试的现代化

《中共中央关于教育体制改革的决定》开宗明义地指出：

“党的十二届三中全会关于经济体制改革的决定，为我国社会生产力的大发展，为我国社会主义物质文明和精神文明的大提高，开辟了广阔的道路。今后事情成败的一个重要关键在于人才，而要解决人才问题，就必须使教育事业在经济发展的基础上有一个大的发展。

教育必须为社会主义建设服务，社会主义建设必须依靠教育。社会主义现代化建设的宏伟任务，要求我们不但必须放手使用和努力提高现有的人才，而且必须极大地提高全党对教育工作的认识，面向现代化、面向未来，为九十年代以至下世纪初叶我国经济和社会的发展，大规模地准备新的能够坚持社会主义方向的各级各类合格人才。”

人才培养中的一个重要问题是人才的考核、鉴定、选拔。我们根据党的教育方针和社会主义现代化建设需要制订各项具体的教学方针、计划、方法，并据此培养出来的各种人才必须经过检验才能判明是否达到要求，这就必须进行教育测量和教育评估。例如，“合格人才”就有质量和数量的要求，不经过测量的手段，就无法了解是否达到要求；同样“各级各类”也要使用测量的手段来进行区分。因此，要提高民族素质，多出人才、出好人才，不但要在教育体制、

教育思想、教育内容、教育方法上进行一系列的改革，还要在考试制度、考试内容和考试方法等方面考虑怎样配合这些改革。教学与考试相辅相成，才能造就人才。

我国是最早建立考试制度的国家之一。西周已建立贡士制度；春秋战国时期，国君、大夫十分注意网罗人才；汉武帝即位后的第一件大事是召集全国文士，亲自命题策问，罢黜百家，独尊儒学。隋唐以后，又建立了逐级选拔的科举制度，开创了以考试方法选拔官吏的先例。科举制度在历史上虽起过一定的积极作用，并积累了不少组织考试的经验；但是因为受到漫长而又日益腐朽的封建制度的影响，科场腐败，考试方法陈旧而又程式化，并没有起到选拔人才的作用，倒是禁锢思想、摧残才智、败坏学风，伤害了人才的成长。

研究怎样运用各种测试手段和统计方法来测量人才的科学叫做教育测量学。尽管在世界各国教育制度中，测试手段已沿用甚久，但是把它们作为一门科学来研究，却还不到一百年的历史。美国哥伦比亚大学教授、著名的教育测量专家 R. Thorndike 认为，教育测量学是在本世纪五十年代初期才进入成熟时期，六、七十年代，由于教育事业发展的需要和电子计算机技术在教育测量中的使用，这门科学有了迅速的发展。在十年浩劫中，我国的教育事业受到疯狂的摧残，考试制度被根本取消，那里谈得上教育测量的研究？1977年，党中央决定恢复高等学校入学的统考制度，马上取得了一连串的积极反应，不但使我国教育事业出现了转机，而且迅速扭转了那种不读书的社会风气。几年来，我国教育事业开始走上了蓬勃发展的道路，各种各样的考试和测验也应运而生，但是由于我们缺乏对教育测量学的基本理论作研究和介绍，在考试中也出现了一些弊端。

§ 1.1 考试中的弊端

一、缺乏一个与教育制度相辅并行的健全的考试制度，教育与考试、人才的测定与人才的选拔存在着脱节现象；国家尚没有管理与组织考试的机构，考试实施无法可循。

二、考试缺乏明确目的，颠倒了教学和考试的关系。一些学校不从根本上去考虑提高教学质量，而是把考试作为推动学习的动力，致使考试成灾，学生穷于应付；又有些学校把主要精力放在帮助学生去对付考试，大搞题海战术，影响了学生的全面发展，削弱了学生的基本功训练；形形色色的应试辅导班、试题集、考试指导书也泛滥成灾，干扰了教学的正常秩序。

三、考试命题缺乏一致的标准和要求。就以我国规模最大的高考而论，命题的范围、重点、深浅程度都不稳定，连考试的要求和方法都经常改变。这就使考试无法对教学起积极的“反拨”作用*。

四、对各种考试的估量缺乏客观标准。对考试的估量应有客观的标准：

(1) 考试是否有效？是否考了应考的内容？是否选拔了应该选拔的考生？

(2) 考试是否可靠？评分是否客观、准确？分数能否反映考生的水平？

(3) 考试能否区分考生的能力？

* 一般来说，考试呼应教学，教学作用于考试；但是一些影响大的考试又往往反过来作用于教学，我们把这种反作用称为“反拨”作用。

(4) 考试是否可行?

但是我们在估量考试时大多是依靠经验性的评论,流于片面。

五、考试的评分的分析方法比较落后。现代世界各国都在使用电子计算机技术和数理统计的分析方法来处理大规模考试的各种数据,以求更好地估量考生的能力。但目前我国还在使用原始分的分数制度,考试中的误差很大,影响到人才的选拔。由于整个考试和录取都使用人力来完成,耗费甚大。

六、不会运用教育测量的方法检验教学试验结果,往往取不到科学的数据,或是取得的数据无效,不足以进行推论。

这些情况都说明:要使我国教育事业适应“四化”的需要,我们必须重视教育测量学的研究,大力改革我国的考试制度、考试内容、考试方法。在实现我国教育事业现代化的过程中,也要实现考试的现代化。考试现代化本身也有“四化”——制度化、专业化、标准化、电脑化,需要我们认真考虑。

§ 1.2 制 度 化

要根据我们党的教育方针和为“四化”培养人才的需要建立符合我国实际的社会主义考试制度。我们需要有中国的考试法,以建立各种发现、培养、审核、选拔人才的制度。考试是一种面向社会的权威的、严肃的、公正的事业,必须实行法治,而不能容许人治的干扰。考试无非是一种量具,关键在于怎样掌握和使用它。使用得当会对教育起到好的促进作用;使用不当,它也象旧时代的科举制度那样营私舞弊、

禁锢思想、败坏学风、贻误人才。考试法正是为了保证考试的公正性和使用得当。用法律形式固定下来的各种考试制度应当有利于我国教育事业的发展，鼓励多种形式、多种层次办学，鼓励自学成才，同时还要规定对考试中各种营私舞弊行为给予应有的处置，乃至法律制裁。根据各国经验，我们应在中央的集中领导下实行考试与使用考试的部门分立的制度，即建立主管考试的机构，专司编制、组织、评定考试的职能，对通过考试的人员发给相应的证书，承认其学历（这是指一些重要的全国性的公开考试）。使用考试的选拔单位或用人单位则应根据德才兼备的原则和工作需要，考虑其考试成绩，量才而用。给予主持考试机构一定的独立地位，有利于保证教育测量的客观性、公正性、权威性，有利于各级教育领导机关掌握教学质量的动态。实施考试的机构主要是建立和主持各种标准化考试和公开考试，作为其他局部性、地区性考试的校正的标尺，以证明通过各种途径成才的人所达到的水平，从而鼓励年青人勤奋学习，积极向上，从风气上和实际上改变那种“一试定终生”的状况。

考试制度必须与教育制度相呼应，使各个层次的教育都有相应的考试证明其学历，而且层次之间还要有所衔接，使自学成才的人能够努力向上，一直至取得最高学历。我国是一个幅员广阔的大国，在统一安排下，发挥中央和地方的两个积极性，共同保证考试制度的实施，至关重要。所谓“制度化”实际上指从教育和考试的相互关系上考虑考试的“反拨”作用，建立一套有利于促进教育发展的法定的考试制度。

§ 1.3 专 业 化

“专业化”有两层意思。一是建立一支教育测量的专业化队伍；二是用教育测量学的专业知识武装各级教育行政管理人员和广大教师。

建立一支教育测量的专业化队伍是为了从组织上保证考试制度的实施。第二次世界大战以后，不少国家和地区都相继成立了专管考试的机构。有的机构根据考试法建立，直属中央政府；有的则是由学校和团体联合资助建立的独立机构。这些机构所主持的考试有的属于国家考试，有的则是标准化考试和公开考试。美国教育考试服务公司是世界上较早成立的一个考试机构，1947年成立时只有24名雇员，现已发展成为拥有2700多名雇员的庞大机构。每年有700多万考生参加它所设计和主持的各种考试。它的考点分布在世界上156个国家。教育考试服务公司所主持的有大型的考试，如专门测量国外留学生英语水平的 TOEFL 考试，有专门测量研究能力的 GRE 研究生考试，有专门测量学习能力的美国大学入学考试 SAT 等；另外还有各种职业性考试；还可根据用人单位的需要临时设计各种考试。教育考试服务公司是一个分工细致、专业性很强的考试机构，它不但设有各种考试的命题和审题委员会（委员中大部分为外单位的专家、教授，少部分为本公司的专职人员），而且设有科研所、计算中心、印刷厂、对外联络办公室、法律顾问处等。这算是大机构的例子。机构小的可以香港考试局为例，该局建于1977年，现亦有专职雇员100多人。该局除主持香港的中学会考和两种大学入学考试外，还代理英美各种考试，有学校考试、

海外考试、行政管理、财务、发展、统计等六个部门。反观我国这么大的一个国家，连一个专司考试的机构还没有成立，这是多么不相适应啊！

建立一个由专业化人员组成的考试机构十分必要，这是因为：

一、考试的组织和试卷的编制是一项保密性强、流程化程度很高的工作，实际上是一项系统工程。其中的每一环节都必须有专人执行、定期完成。以 TOEFL 考试为例，从决策到最后拼成试卷约需经过11个大的环节（不包括报名、考试组织等行政事务工作），牵涉到五、六十人，而这个考试每月都要举行一次，这就意味着命题组每月都要出一套试卷。没有专人层层负责，实行作业流程化，这几乎是不可能的。我国目前不一定需要编制这样的试卷，但是根据标准化考试的要求，出一套试卷，前后约需一年时间（起码也要半年），才能保证质量。在这么长的一段时间内，没有专人按部就班地掌握每个环节，也是很困难的。

二、考试的试卷要有继承性和均衡性。有些定期举行的考试要从内容、深浅程度、题型、考试要求等方面保持一致性。如果一个考生水平没有变化，他在不同时期内所参加的同一个考试的成绩应该一致。靠临时召集一些人来命题，就会有损试卷的标准性，无法提供一把稳定的尺子，也难以使考试对教学起到好的推动作用。命题组必须有一部分稳定的命题委员，而考试机构里的专业人员担任稳定的命题人员正可以保证试题的继承性和均衡性。

三、考试的各种资料必须集中管理和保存，才能起作用。在一些考试机构里常设有题库，把经过试验，证明可靠性高、区别性强的题目按难易程度、题型、内容编成卡片资

料保存，供拼题时使用。题库必须由专人建立和保管，才能发挥作用，因为它有很高的保密要求。

四、考试的成绩必须集中统计分析，才能掌握教学质量、考生水平的全面情况，开展横向和纵向研究，为教育领导部门决策时提供参考数据。

五、对教育测量开展科学研究。考试是牵动千家万户的大事，考试的改革要坚持科研先行；只有经过试验证明为有效的改革才能逐步引进到大规模的考试里。此外，还有许多有待探索的问题，需要有专门人员去研究。

六、和国外考试机构建立联系，交换考试情报。专业化人员应该包括三支队伍：一是命题人员；二是统计分析人员；三是考试的组织和管理人员。

命题人员不可能都是脱产的，但是试题的编写又以他们为主体，这就产生另一个问题：用现代教育测量学的理论和技巧武装命题人员。命题人员应该既是某一学科的专家，同时又比较熟悉测试的理论和方法，才能出好题目。在教学行政管理人员和教师中也应普及教育测量学的常识，他们是考试的使用者，亦应了解考试所提供的反馈信息，以改进教学。在教育学院、师范学院、教师进修学院应普遍开设教育测量学的课程，或举办短训班；还要组织力量编出若干种反映国外先进水平和结合我国实际的教育测量学教材。

§ 1.4 标 准 化

“不以规矩，不能成方圆”。在实现现代化过程中，不可以没有标准化。计量的标准化对组织工农业生产和开展科学研究尤为必要。教育测量的对象是人，而人的因素却是变